

LINGUISTIC DICTIONARY AND METHOD FOR PRODUCTION THEREOF

ABSTRACT OF THE DISCLOSURE

A method and arrangement for handling case and other orthographic variations in linguistic databases by explicit representation comprising: explicit storage of all orthographic and case variations of words in the dictionary, and use of extended cut and paste codes to control dictionary size explosion and to make the restoration of the lemma more efficient. This provides the advantage of allowing very efficient handling of case and orthographic variants while performing a dictionary lookup.